

Die Rezeption (Thyssen-)preisgekrönter Artikel in der „Scientific Community“

Andreas Diekmann · Matthias Näf · Manuel Schubiger

Zusammenfassung: Wir untersuchen in unserer Studie, ob Artikel, die mit dem Thyssen-Preis prämiert wurden, in höherem Maße rezipiert werden als nicht-ausgezeichnete Artikel. Dazu wurden die Ergebnisse für die prämierten Artikel mit einer Kontrollstichprobe der nicht ausgezeichneten Arbeiten verglichen. Prämierte Artikel „ernteten“ signifikant mehr Zitationen als Artikel in der Kontrollstichprobe. Bemerkenswert ist, dass das Muster der Zitationsrangfolge exakt der Rangfolge der Preise entspricht. Der erste Preis erhält die meisten, der zweite Preis die zweitmeisten und der dritte Preis die drittmeisten Zitationen, während nicht prämierte Artikel an vierter Stelle der Rangfolge stehen. Der Unterschied zeigt sich auch, wenn nur Zitationen im Jahr der Veröffentlichung eines Artikels und im Folgejahr betrachtet werden. In diesem Zeitraum kann der Preis „an sich“ noch keine Wirkung auf die Rezeption in der „scientific community“ entfaltet haben. Die Ergebnisse liefern keinen Beleg für die sozialkonstruktivistische These, dass die Rezeption im Sinne einer „sich selbst erfüllenden Prognose“ erst durch die Preisverleihung erzeugt wird. Vielmehr gelingt es der Jury Arbeiten auszuwählen, die dann tatsächlich in den Folgejahren in der Fachöffentlichkeit Aufmerksamkeit hervorrufen.

Schlüsselwörter: Wissenschaftssoziologie · Wissenschaftspreise · Impact von Artikeln · Matthäus-Effekt · Thomas-Theorem · Bibliometrie

The impact of (Thyssen)-awarded articles in the Scientific Community

Abstract: We investigate whether journal articles winning a prestigious award of the Thyssen foundation have a higher impact in the scientific literature than a control group of non-awarded

© VS Verlag für Sozialwissenschaften 2012

A. Diekmann (✉) · M. Näf
ETH Zürich, CLU D3,
Clausiusstrasse 50, 8092 Zürich, Schweiz
E-Mail: andreas.diekmann@soz.gess.ethz.ch

M. Näf
E-Mail: matthias.naef@soz.gess.ethz.ch

M. Schubiger
ETH Zürich, 8092 Zürich, Schweiz
E-mail: schumanu@ee.ethz.ch

articles. In average, awarded articles harvest significantly more citations than articles in the control group. Most remarkably, the average citation rank exactly matches the rank order of awards. The top award earns most citations while the second award, the third award and the non-awarded articles exhibit citation counts in declining order. The correlation of award rank and impact does not vanish if we restrict citation counts to the year of publication and the year thereafter. In this time span the distinction of excellence is very unlikely having a genuine effect on the impact measure. Thus, we do not find support for the social constructivist hypothesis that the correlation of the jury decision and the article impact is explainable by a label of distinctiveness. On the contrary, the empirical data confirm that the jury is successful in the selection of articles which in average have a larger impact in the scientific community than non-awarded publications.

Keywords: Sociology of scientific knowledge · Scientific awards · Impact of articles · Matthew effect · Thomas theorem · Bibliometrics

1 Einleitung

In der Evaluationsforschung werden quantitative Indikatoren ebenso wie qualitative Expertenurteile herangezogen. Oft wird auch zweigleisig gefahren, wie etwa beim Forschungsrating Soziologie des Wissenschaftsrats (Neidhardt 2008). Bei diesem Pilotprojekt wurden neben der Erhebung zahlreicher Leistungsmerkmale der beteiligten Institutionen auch Expertenurteile zur Qualität der Forschungsleistung berücksichtigt. Es stellt sich die Frage, ob qualitative Expertenurteile und quantitative Indikatoren zu ähnlichen oder gar zu divergierenden Einschätzungen führen. Dieser Frage wird anhand einer systematischen Analyse der Zitationen von Artikeln nachgegangen, die mit dem Thyssen-Preis gewürdigt wurden. Aufschlussreich ist hierbei insbesondere der Vergleich mit einer Kontrollgruppe aus den Publikationen der Zeitschriften, die regelmäßig am Thyssen-Preis teilnehmen. Einen solchen Vergleich könnte man auch als Validierung eines Expertenurteils anhand eines Außenkriteriums bezeichnen. Ein positiver Zusammenhang zwischen Expertenurteil und Rezeption in der wissenschaftlichen Literatur könnte bedeuten, dass die Jury-Mitglieder, wie man erwarten möchte, ein Gespür für die Qualität wissenschaftlicher Artikel haben, die dann in der Forschergemeinschaft auch auf entsprechende Resonanz stoßen. Es gibt aber auch die alternative, „konstruktivistische“ Erklärung, dass im Sinne des „Thomas-Theorems“ mit der Preisvergabe eine Situation definiert wird, deren Konsequenzen real sind (Thomas und Thomas 1928).¹ Ein Artikel mit dem Prädikat „Thyssen-Preis“ würde dann nicht aus Gründen der Qualität, sondern aufgrund der Wahrnehmung der Auszeichnung häufiger zitiert als ein „normaler“ Artikel. Zu bemerken ist weiterhin, dass die Kontrollstichprobe einen eigenen Wert hat, der über die Ziele der vorliegenden Studie hinausgeht. Man erhält mit den Zitationen für die

1 Das bekannte Zitat („If men define situations as real, they are real in their consequences“, S. 572) wird eher beiläufig in der Studie von William und Dorothy Thomas über „The Child in America“ erwähnt. In ihrem auch heute noch lesenswerten Kapitel über „The Methodology of Behavior Study“, in dem sich das berühmte Zitat findet, wird u. a. dafür plädiert, die Einflüsse verschiedener Situationsmerkmale auf das Verhalten mit statistischen Methoden zu untersuchen und dabei einzelne Merkmale schrittweise genauer zu kontrollieren. Bekannt wurde das Zitat durch andere Arbeiten, insbesondere durch diverse Artikel von Robert K. Merton. Siehe zur Geschichte des Thomas-Theorems Merton (1995).

Kontrollstichprobe einen Referenzwert für die Häufigkeit, mit der ein deutschsprachiger Fachartikel in referierten Zeitschriften im Durchschnitt zitiert wird. Damit wird auch ein Ausgangspunkt für internationale Vergleiche festgelegt. Denn alle Unkenrufe über eine geringe Rezeption deutschsprachiger sozialwissenschaftlicher Artikel beruhen nur auf Mutmaßungen.

Kürzlich haben Alber und Fliegner (2010) sowie Alber et al. (2010) Befunde aus einer breit angelegten Untersuchung über Autoren, Themen, Rezeption und weitere Merkmale von Aufsätzen berichtet, die mit dem Thyssen-Preis ausgezeichnet wurden. Eine Analyse der Rezeption der „Thyssen-Aufsätze“ anhand des „Social Science Citation Index“ (SSCI) erfordert allerdings eine Vergleichsgruppe. Nur so kann eine Aussage darüber getroffen werden, ob die prämierten Aufsätze häufiger als andere Artikel rezipiert wurden. Parallel zu unseren Auswertungen hat Jansen (2012) in einer aufwendigen und bemerkenswerten Arbeit die mit dem Thyssen-Preis gewürdigten, die nominierten, aber nicht ausgezeichneten und die nicht nominierten Artikel bezüglich der Häufigkeit von Google-Scholar Zitationen verglichen. Im Einklang mit den Erwartungen hat er deutliche und signifikante Unterschiede der Zitationshäufigkeiten gefunden. Wir verwenden hier die SSCI-Zitationen, die über das Echo auf eine Publikation in der „Scientific community“ Aufschluss geben. Die Artikel wurden manuell verkodet, da nicht alle am Thyssen-Preis teilnehmenden Zeitschriften im SSCI geführt werden.² Eine wesentliche Innovation der vorliegenden Arbeit ist die Herstellung einer Vergleichsstichprobe. Zu diesem Zweck wurde eine Zufallsstichprobe aus sämtlichen 6625 Artikeln gezogen, die im Zeitraum von 1981 bis 2008 in den am Auswahlverfahren für den Preis beteiligten Zeitschriften veröffentlicht wurden. Darüber hinaus gehen wir einen Schritt weiter und versuchen Licht auf das Kausalitätsproblem zu werfen. Finden wir die erwarteten Unterschiede in der Rezeptionshäufigkeit, so stellt sich die Frage, ob die Jury als Filter fungiert oder eventuell bestehende Unterschiede durch eine Wirkung der Jury als „Etikettierungsinstanz“ erklärbar sind.

2 Rezeptionserfolg von Artikeln

„What makes an article influential?“ fragen Sternberg und Gordeeva (1996). Dieser Frage wurde mit einer Erhebung unter Mitgliedern der US-amerikanischen Psychologengemeinschaft nachgegangen. Herausgekommen sind sechs Dimensionen: 1) die Qualität der Präsentation, 2) die theoretische Signifikanz, 3) Implikationen für die Praxis, 4) die Bedeutung und Aktualität der Thematik, 5) methodische Innovationen und 6) die Impli-

2 Für jeden einzelnen Artikel werden bei diesem Verfahren die Zitationen im SSCI ermittelt. In der von Jansen (2010) eingerichteten Datenbank (unsere Erhebung hat im Frühjahr 2010 begonnen und wir haben von der Datenbank leider erst nach Abschluss unserer Arbeiten erfahren) wird nur den in „ISI-Zeitschriften“ publizierten Artikeln ein Zitationswert zugeschrieben. Mit dieser Teilmenge würde man aber systematisch verzerrte Werte erhalten. Alber und Fliegner (2010) haben ebenfalls den prämierten Aufsätzen SSCI-Zitationen manuell zugewiesen. Sie haben aber keine Kontrollgruppe gebildet und auch nicht nach Art des Preises unterschieden. Siehe auch weiter unten zur Methode der Untersuchung.

kationen eines Artikels für die künftige Forschung. Dies sind die normativen Vorstellungen über die Qualität eines „High-Impact-Artikels“.

Wie wir aus den klassischen Arbeiten von Merton (1973) wissen, sollte das Unternehmen „Wissenschaft“ einem universalistischen Programm folgen. Wissenschaftliche Erkenntnis ist unpersönlich und wird an *pre established impersonal criteria* gemessen, „in Übereinstimmung mit den Beobachtungen und vorgängig bestätigtem Wissen. Die Annahme oder Ablehnung von wissenschaftlichen Behauptungen (...) soll nicht von persönlichen oder sozialen Eigenschaften ihrer Protagonisten abhängen. Ihre Rasse, Nationalität, Religion, soziale Klasse und persönliche Eigenschaften sind als solche irrelevant. Objektivität verbietet Partikularismus“ (Merton 1973, S.270). Und weiter heißt es: Ein „anglophober Forscher kann nicht das Gravitationsgesetz außer Kraft setzen. (...) Der Imperativ des Universalismus ist tief verwurzelt im unpersönlichen Charakter der Wissenschaft.“

Entsprechend sollte nach der universalistischen Idealvorstellung die Qualität eines Fachartikels über den Grad der Rezeption in der wissenschaftlichen Zunft entscheiden. Anonyme Peer-Begutachtungen werden arrangiert, um partikularistische, sachfremde Aspekte bei der Entscheidung über die Publikation eines Manuskripts auszublenden.

Der universalistischen, normativen Auffassung von Wissenschaft kann die These des Sozialkonstruktivismus gegenübergestellt werden (Baldi 1998; Bornmann und Daniel 2008). Danach haben Macht, Status, Rhetorik und andere partikularistische Merkmale einen entscheidenden Einfluss auf die Rezeption und Akzeptanz wissenschaftlicher Erkenntnisse. Diesen Einflüssen wird bevorzugt in Untersuchungen zu Wissenschaftskulturen nachgegangen, in denen Wissenschaftlern im Labor gewissermaßen über die Schulter geschaut wird. So heißt es bei Knorr-Cetina (1989, S. 91): „Das empirische Programm des Konstruktivismus stellt zunächst den Versuch dar, die jeweils involvierten Konstruktionsprozesse, die eigentliche Konstruktionsmaschinerie zum Gegenstand der Analyse zu machen“. Auch Merton (1968) weiß, dass Verbreitung und Akzeptanz wissenschaftlicher Erkenntnisse nicht nur von der Qualität und objektiven Beweiskraft der Argumente und Beobachtungen abhängen. Seine berühmte Arbeit über den „Matthäus-Effekt“ betont ja gerade, dass das „außerwissenschaftliche“ Merkmal der Prominenz eines Autors Aufmerksamkeit und Verbreitung von Forschungsergebnissen befördert. Dafür gibt es ganz rationale Gründe. Zeitknappe Leser achten auf die Reputation, da ein für herausragende Leistungen bekannter Forscher mit hoher Wahrscheinlichkeit wieder Qualitätsprodukte erzeugen wird. Der Matthäus-Effekt, so die Kommunikationshypothese von Merton (1968), führt zur beschleunigten Verbreitung von Forschungsergebnissen. Die Folgen sind positiv für das Wissenschaftssystem, wenn auch oft nachteilig für einzelne Forscher, die mit den gleichen Entdeckungen oder anderen originellen Beiträgen im Schatten ihrer prominenten Kollegen stehen. Merton erkennt aber auch die Nachteile des Effekts, wenn nämlich Autoritätshörigkeit zu dogmatischer Verkrustung führt.

Der Matthäus-Effekt basiert auf Reputation. Ebenso kann ein Preis Reputation und Aufmerksamkeit verschaffen. Wenn ein preisgekrönter Artikel häufiger zitiert wird, könnte es der Fall sein, dass die erhöhte Rezeption nicht der Qualität zu verdanken ist (und damit das Jury-Urteil rechtfertigt), sondern einem preisgekrönten Artikel einfach mehr Aufmerksamkeit zuteil wird.

Welche Faktoren sind aus empirischer Sicht verantwortlich für den mehr oder minder großen Einfluss eines Artikels auf die Fachdiskussion? Haslam et al. (2008) beklagen, „there has been relatively little research on predictors of article impact“. Immerhin liegen aber eine Reihe von Arbeiten vor, die sich auf empirischer Grundlage mit dem Einfluss von Merkmalen des Autors (Geschlecht, Berühmtheit), dem Prestige der Institution, dem Prestige des Journals und anderen Faktoren befassen, die womöglich einen Einfluss auf den „Impact“ eines Artikels, gemessen an der Häufigkeit von Zitationen, ausüben.

In der Studie von Haslam et al. (2008) sind die folgenden Faktoren bedeutsam für die Zitierhäufigkeit von Artikeln in Psychologie-Fachzeitschriften: Die Prominenz eines Autors, das Prestige des Journals, die Länge eines Artikels, die Anzahl und Aktualität der im Artikel zitierten Literatur; das Geschlecht des Erstautors spielte dagegen keine Rolle. Judge et al. (2007) unterscheiden „universalistische“, „partikularistische“ und „gemischte“ Prädiktoren.

Zu ersteren zählen die Originalität, die Methodik der Gewinnung und die Präsentation („writing“) von Forschungsergebnissen. Partikularistisch sind Merkmale wie die Prominenz eines Autors, das Prestige der Institution und das Geschlecht. Zu der Restkategorie zählen das Prestige des Journals und die Platzierung des Artikels in der Zeitschrift. So hatten Gespräche mit Herausgebern ergeben, dass die wahrgenommene Qualität insbesondere für die Entscheidung über den erstplatzierten „Lead“-Artikel (die „pole position“) ausschlaggebend sei. Judge et al. (2007) untersuchen Artikel aus der Management-Literatur. Sie finden, nicht ganz überraschend, heraus, dass das Prestige eines Journals für die Rezeption von herausragender Bedeutung ist. Wichtig ist ferner das Prestige der Institution bei theoretischen Artikeln; keinen oder einen sehr geringen Einfluss hat die Institution dagegen bei empirisch orientierten Artikeln. Das Geschlecht des Erstautors spielte keine Rolle. Dass Meta-Analysen sehr viel häufiger zitiert werden als andere theoretische Artikel, liegt auf der Hand. Auch in dieser Arbeit zeigt sich wieder ein relativ starker Effekt der Anzahl im Literaturverzeichnis aufgeführter Referenzen. Eine mögliche Erklärung ist Reziprozität. Eine Zitation ist ein freundlicher Akt, der mit Freundlichkeit erwidert wird (Judge et al. 2007). Die Autoren deuten ihre Ergebnisse insgesamt als Hinweis auf „rather small effects of particularism on citations“ (Judge et al. 2007, S. 503).

Im Grunde sagen die Ergebnisse dieser Studien aber wenig über universalistische oder partikularistische Effekte aus, da sich der Einfluss dieser Merkmale kaum in die erwähnten Kategorien zerlegen lässt. Wenn der Artikel eines prominenten Autors häufiger zitiert wird als ein unbekannter Autor, kann dies am „Matthäus-Effekt“ liegen, es kann aber auch ganz einfach der Fall sein, dass ein erfolgreicher und prominenter Autor bessere Qualität liefert. Gleiches gilt für das Prestige einer Institution und insbesondere für das Prestige eines Journals.

Einen methodisch anderen Ansatz wählte Baldi (1998) in einer Untersuchung über Zitierhäufigkeit in der Astrophysik. Sein „Netzwerk-analytisches Modell“ betrachtet jeweils Dyaden von Artikeln. Analyseeinheit ist die Dyade von (potenziell) zitiertem und (potenziell) zitierendem Artikel. Merkmale beider Glieder einer Dyade fließen in die Analyse ein, wobei die Qualität eines Artikels über die Zitationsrate gemessen wird. Theoretischer Ausgangspunkt ist die Kontroverse zwischen „normativem“ Ansatz und sozialkonstruktivistischer Perspektive. Bei ersterem gilt: „that citations reflect the worth and content of contributions“. Die sozialkonstruktivistische Position lautet dagegen: „...

that citations are allocated largely on the basis of an author's position within the stratification structure of science". Oder prägnanter auf die Kurzformel gebracht: „what one says" versus „who one is" (Baldi 1998, S. 835). Wird für Merkmale der Dyade kontrolliert, zeigt sich kein Effekt der sozialkonstruktivistischen Merkmale. Auch die Prominenz eines Autors spielt für die Zitationen in der Astrophysik in Baldi's Studie keine Rolle; einzig das Geschlecht eines Autors hat einen Effekt, und zwar zu Ungunsten der Physikerinnen.

Können Gutachter den Rezeptionserfolg eines Artikels vorhersagen? Gottfredson (1978) hat Expertenurteile zu mehr als 500 in psychologischen Fachzeitschriften veröffentlichten Artikeln erhoben. U. a. sollte die Rezeption der Artikel eingeschätzt werden. Der Zusammenhang zwischen Expertenurteil und Rezeptionserfolg, gemessen wiederum über die Zitationsrate, war mit einem Korrelationskoeffizienten von 0,24 (bei Durchschnittsbildung mehrerer Expertenurteile zum gleichen Artikel $r=0,27$) nur mäßig hoch. Preisverleihungen sind eine Gelegenheit, um den Zusammenhang zwischen Expertenurteilen und Rezeption von Artikeln nicht-reaktiv mit prozessproduzierten Daten zu untersuchen. Eine derartige Studie wurde von Lee et al. (2003) durchgeführt. Grundlage der empirischen Arbeit waren 1682 Artikel auf dem Spezialgebiet ergonomischer Forschung, die 1970 bis 2000 im führenden Journal des Fachgebiets „Human Factors" veröffentlicht wurden. Im Zeitraum der Untersuchung wurden 30 Preise vergeben. Zwischen dem Rezeptionserfolg der preisgekrönten und den nicht-ausgezeichneten Artikeln besteht ein deutlicher und signifikanter Unterschied. Die ausgezeichneten Artikel wurden im Durchschnitt 1,32 Mal pro Jahr zitiert, die übrigen Artikel 0,59 Mal. Lee et al. stellen auch die Frage nach der Kausalität, beantworten sie aber nicht auf der Grundlage der eigenen Daten. Könnte es, wie oben bereits angedeutet, sein, dass die Etikettierung eines Artikels durch die Preisverleihung, also ein Akt sozialer Konstruktion, erst die scheinbare Qualität schafft, die mit der Auszeichnung suggeriert wird? Neben der Frage nach Richtung und Stärke des Zusammenhangs zwischen Thyssen-Preis und Zitationsrate wird uns auch interessieren, ob eine zu erwartende Korrelation womöglich erst durch die „Etikettierung" hergestellt wird. Man könnte sogar aus sozialkonstruktivistischer Perspektive darüber spekulieren, dass die Etikettierung einen doppelten Effekt auslöst. Eine Preisverleihung verleiht Aufmerksamkeit und erhöht die Chance der Anhäufung von Zitationen. Wer häufig zitiert wird, erzielt wiederum mehr Aufmerksamkeit. Ein dem Matthäus-Effekt analoger Mechanismus löst eine sich selbst verstärkende Zitationsspirale aus. Aus der Hypothese folgt, dass zwei Artikel A und B gleicher fachlicher Qualität stark unterschiedliche Häufigkeiten von Zitationen „ernten" können, wenn Artikel A prämiert wird, Artikel B hingegen keine Auszeichnung erhält.

Einen empirischen Beleg zur Validität von Expertenurteilen liefert die Studie von Bornmann und Daniel (2006) über eine Preisvergabe an Naturwissenschaftler. Der Fonds des Pharma-Unternehmens Boehringer-Ingelheim vergibt per Jury-Entscheid ein Post-Doc-Stipendium an hoch qualifizierte Bewerber. Verglichen wurden die Zitationsraten von akzeptierten und abgewiesenen Bewerbern für Arbeiten, die vor der Jury-Entscheidung publiziert wurden. Diese waren bei den akzeptierten Kandidaten deutlich höher als bei den abgewiesenen Bewerbern. Die Preisvergabe selbst konnte also keinen Effekt auf die Zitationsrate ausgeübt haben, sodass der konstruktivistische Einwand hier nicht stichhaltig ist. Allerdings hatten die Jury-Mitglieder Kenntnis von den Publikationslisten der Bewerber. Wenn in das Jury-Votum auch nicht direkt die Zitationsraten eingeflossen sind,

so hatten die Experten doch ein Bild von der Publikationsleistung und insbesondere dem Prestige der Zeitschriften, in denen die Bewerber publiziert hatten. Zwar wurde bei dem Vergleich für die Häufigkeit der Publikationen kontrolliert (es wurden durchschnittliche Zitationsraten pro Artikel berechnet). Jedoch kann nicht ausgeschlossen werden, dass das Jury-Votum per Verfahren indirekt mit den Zitationsraten korreliert ist. Das Renommee der Fachzeitschriften ist ein „Drittfaktor“, der mutmaßlich sowohl mit dem Urteil von Juroren und Gutachtern als auch mit der Zitationsrate korreliert ist.

Beim Thyssen-Preis entscheiden die Herausgeber der Zeitschriften und die Jury dagegen nicht über eine Person mit vorliegender Publikationsliste, sondern über einen Artikel. Ein weiterer Vorteil für unsere Studie besteht darin, dass für jeden Zeitschriftenjahrgang ein erster, zweiter und dritter Preis vergeben wird. Dadurch erhöht sich nicht nur die Zahl der Beobachtungen. Es besteht auch die Möglichkeit, die Rezeption der Artikel separat nach der abgestuften Klassifizierung der Jury zu untersuchen.

3 Zur Methodik der Untersuchung

Der Preis der Fritz-Thyssen-Stiftung für sozialwissenschaftliche Aufsätze wird seit dem Jahr 1981 jährlich für die, nach dem Urteil einer Jury von Experten, besten Aufsätze vergeben, die innerhalb eines Jahres in den teilnehmenden Fachzeitschriften publiziert wurden (Liste der Zeitschriften im Online-Anhang 1). Das Verfahren ist zweistufig: Zunächst nominieren die Herausgeber der Zeitschriften bis zu zwei Artikel. Die ausgewählten Artikel werden sodann von einer Jury bewertet, die einen ersten, zweiten und dritten Preis vergibt.³ Aufgrund der Teilung von Preisen kommt es vor, dass oftmals mehr als drei Artikel pro Jahr prämiert werden.

Berücksichtigt werden in unserer Untersuchung alle 102 im Zeitraum 1981–2008 prämierten Artikel (Liste der Artikel und ihrer Zitationshäufigkeiten im Online-Anhang 1). Zusätzlich wurde eine Kontrollstichprobe von 146 Artikeln gezogen (siehe Online-Anhang 1).

Wie zieht man eine Kontrollstichprobe von Artikeln? Man könnte z. B. von jeder Zeitschrift in jedem Jahr der Beteiligung am Preisverfahren das zweite Heft eines Jahrgangs und aus diesem einen Artikel auswählen. Allerdings müsste man dann noch berücksichtigen, dass die einzelnen Fachzeitschriften sich nach der Zahl der jährlich publizierten Artikel und dem Teilnahmezeitraum unterscheiden. Wir haben uns stattdessen für das Verfahren einer einstufigen Zufallsauswahl entschieden. Sämtliche Artikel der 17 Zeitschriften, die potenziell eine Chance der Nominierung im Zeitraum seit der ersten Preisvergabe bis zum Jahre 2008 hatten, wurden (elektronisch) unter Verwendung der Datenbank SOLIS (und teilweise manuell) durchnummeriert. Das ergab eine Grundgesamtheit von 6625 Artikeln. Aus dieser Population von Fachzeitschriftenaufsätzen mit Nominierungschance wurde per Zufall eine Stichprobe von 150 Artikeln gezogen. Vier Artikel ent-

³ Auch Mitglieder der Jury können direkt Artikelvorschläge einbringen. Die Jury besteht derzeit aus acht Mitgliedern, die kooptiert wurden. Für Auskünfte bedanken wir uns bei Heiner Meulmann. Zur Zusammensetzung der Jury siehe auch die Web-Seite zum Thyssen-Preis: <http://www.fis.uni-koeln.de/3121.html>.

sprachen nicht den Kriterien für eine Nominierung (ein Nachruf und drei Kommentare), sodass nach Ausschluss dieser Artikel eine Zufallsstichprobe von 146 Artikeln vorlag.

Für sämtliche Artikel beider Stichproben (d.h. einer Stichprobe und einer Vollerhebung) wurde die Häufigkeit der Zitationen im „ISI Web of Knowledge“ ermittelt. Genauer: Die Zitationen im SSCI (plus Science Citation Index SCI, Art & Humanities Citation Index (A&HCI) und die anderen Datenquellen des Web of Science). Der Einfachheit halber sprechen wir im Folgenden von SSCI oder ISI. Warum „ISI“ und nicht z. B. Google-Scholar?

Google-Scholar ist eine Suchmaschine für wissenschaftliche Literatur. Sie macht keinen Unterschied zwischen Zitaten in studentischen Seminararbeiten, Hinweisen auf Literatur in Veranstaltungsprogrammen, Zitaten in Arbeitspapieren oder aber in publizierten Fachartikeln. Ein Lehrbuch, das in der Ausbildung beliebt ist, erzielt auch entsprechend viele Zitate in Google-Scholar. Zuviel Spreu und weißes Rauschen in den Ergebnissen der Suchmaschine erlauben womöglich nur eine grobe Abschätzung des Rezeptionserfolgs einer Veröffentlichung in der „Scientific Community“. Für das Ausmaß der Rezeption von Artikeln in Fachzeitschriften könnte aber auch Google Scholar zu validen Resultaten führen. Ob Google-Scholar näherungsweise valide Resultate hervorbringt, ließe sich ermitteln, indem der Indikator mit anderen bibliometrischen Indikatoren korreliert wird.

Im SSCI dagegen werden die Zitationen in sämtlichen Zeitschriften erfasst, die dem ISI-Club angehören. Das sind in der Regel referierte und zumeist, aber nicht nur, englischsprachige Fachzeitschriften. Wie man auch in Expertengremien immer wieder erfahren darf, gibt es erhebliche Missverständnisse über die Möglichkeiten der Datenbank. Dazu kurz einige Anmerkungen:

- Die Anzahl der Zitate ist noch kein Gradmesser der Qualität einer Veröffentlichung. Ein Forschungsergebnis, das häufig zitiert wird, kann sich später als falsch (vielleicht sogar als gefälscht?) herausstellen, die Wissenschaft kann einer Modeströmung folgen oder einem „kollektiven Irrtum“ verfallen u. a. m. Außerdem variiert die Zitationshäufigkeit enorm nach Disziplinen. Dagegen ist die Zitationshäufigkeit eine Messlatte für die Aufmerksamkeit, die eine Veröffentlichung in der Wissenschaftsgemeinschaft erzielt. Die SSCI-Werte sind ein Indikator für das Ausmaß der Rezeption („impact“) einer Publikation in der Zunft der in (ISI-)Fachzeitschriften schreibenden Wissenschaftlerinnen und Wissenschaftler. Womöglich korrelieren Qualität und Rezeption, wobei aber die Antwort auf diese Frage für unsere Studie keine Rolle spielt.
- Der SSCI ist englischsprachig dominiert. Das ist richtig und falsch zugleich. Natürlich ist Englisch heute in wachsendem Maße die Sprache der Wissenschaft. Es ist aber keineswegs so, dass eine soziologische Veröffentlichung in einer englischsprachigen Fachzeitschrift zwangsläufig der Sprache wegen mehr Zitate erhält. Immerhin neun der am Thyssen-Preis beteiligten, deutschsprachigen Fachzeitschriften sind zugleich im SSCI aufgeführt. Einige deutschsprachige Zeitschriften haben höhere Impact-Faktoren als eine durchschnittliche englischsprachige Spartenzeitschrift. Der Grund dafür ist, dass der deutschsprachige (soziologische) Wissenschaftsraum relativ groß und eine Welt für sich ist. Ein Artikel in der Kölner Zeitschrift für Soziologie und Sozialpsychologie wird z. B. von einem Autor in der Sozialen Welt, dem Berliner Journal und der Zeitschrift für Soziologie zitiert und der SSCI verzeichnet drei zusätzliche

Zitate. Kontroversen innerhalb der deutschen Soziologie finden ihren Niederschlag im SSCI, ohne dass die Kontroverse jenseits der Grenzen Wellen schlagen muss. Deshalb ist der SSCI auch nur begrenzt ein Maß für die Internationalität der Rezeption. (Dafür müsste man nach Zitaten in deutschsprachigen Zeitschriften und Zitaten in Zeitschriften unterscheiden, die außerhalb des deutschen Sprachraums erscheinen.)

- Ein weiteres Missverständnis lautet, nur Zitate aus einer kleinen Auswahl von Fachzeitschriften würden erfasst. Verwechselt werden dabei zitierte Quelle und zitierende Zeitschrift, also quasi Sender und Empfänger. Bei den zitierenden Zeitschriften kommen nur ISI-registrierte Zeitschriften in Frage. Als zitierte Quelle kann dagegen jeder Text (Monographien, Sammelbände, alle Arten von Fachzeitschriften, unveröffentlichte Arbeitspapiere usw.) Berücksichtigung finden. Im SSCI finden wir demnach auch die Zitationen der am Thyssen-Preis beteiligten Zeitschriften, die nicht im SSCI registriert sind. (Nebenbei bemerkt ist der SSCI ein sehr gutes Instrument, um die Rezeption von Büchern in der Fachgemeinschaft zu ermitteln. Die Verteilung ist vermutlich extrem rechtsschief (linkssteil). Es gibt deutschsprachige Bücher, die eine drei- oder in Einzelfällen eine vierstellige Anzahl von Zitaten auf sich vereinigen, während ein Großteil der Monographien niemals oder sehr selten zitiert wird.) Allerdings kann man die Zitate externer Quellen nicht einfach auf Knopfdruck ermitteln.⁴ Wenn man jedoch, wie in unserem Falle, die Quelle kennt, kann man die Zahl der Zitationen eines Artikels mit ziemlicher Genauigkeit bestimmen.

Dafür wird jeder Artikel aus der Kontrollgruppe und der Gruppe der prämierten Artikel separat in das System eingegeben. Wenn mindestens zwei der drei Merkmale der zitierten Quelle, die im SSCI angezeigt werden (Erscheinungsjahr, Band (Volume), Seitenzahl) mit der im zitierenden Artikel genannten Quelle übereinstimmen, wird das Zitat gezählt. (Einige wenige Zitate bleiben dadurch unberücksichtigt, z. B. wenn ein Autor eine falsche Seite und eine falsche Nummer des Zeitschriftenbands angibt.) Alle Eingaben wurden sodann von einer zweiten Person bei einem weiteren Durchgang überprüft. Dabei wurden nur wenige Abweichungen festgestellt. Die Ergebnisse basieren auf den Daten der wiederholten Kodierung, wobei Abweichungen gegenüber der ersten Kodierung korrigiert wurden.⁵

Während die Artikel in der Kontrollstichprobe und in der Gruppe der prämierten Aufsätze im Zeitraum 1981–2008 publiziert wurden, werden die Zitationen über diesen Zeitraum hinaus bis Ende 2011 erfasst. Um ferner Effekte des Erscheinungsjahrs zu kontrollieren (ein 1981 publizierter Artikel hat natürlich eine größere Chance, bis 2011 viele Zitate zu sammeln, als ein Artikel, der 2005 publiziert wurde) standardisieren wir nach Fünf- und nach Zehn-Jahres-Zeitintervallen. Nur Artikel, für die ein vollständiges „Beobachtungsfenster“ vorliegt, werden in die Analyse einbezogen. Beispielsweise erstreckt

4 Auf „Knopfdruck“ kann man dagegen leicht ermitteln, wie häufig ISI-Artikel von ISI-Artikeln zitiert werden. In der Soziologie, in der das Medium Buch eine große Rolle spielt, führt dieses „Quick-and-dirty“-Verfahren dagegen zu einer systematischen Verzerrung.

5 Genau genommen gab es noch einen dritten Durchgang, bei dem, der Anregung eines Gutachters folgend, zwischen Selbst- und Fremdzitaten unterschieden wurde. Dabei konnte gleichzeitig der Zitationszeitraum (ursprünglich bis Frühjahr 2010) bis Ende 2011 aktualisiert werden.

sich der Beobachtungszeitraum für einen 2007 publizierten Artikel über vier nachfolgende Jahre (bis Ende 2011), sodass dieser Artikel bei der Ermittlung der Zitationen für den Fünf-Jahres-Zeitraum unberücksichtigt bleibt. Entsprechend sinkt die Fallzahl mit zunehmender Länge des Beobachtungsintervalls.

4 Die Rezeption prämierter Aufsätze im Vergleich zur Kontrollstichprobe

Wir stützen unsere Aussagen auf Vergleichswerte für standardisierte Zeitspannen. Nach fünf Jahren wird ein prämierter Artikel 3,3 Mal zitiert, ein Artikel in der Kontrollstichprobe im Durchschnitt nur wenig mehr als ein Mal (1,1). Nach zehn Jahren betragen die entsprechenden Zahlen 6,1 versus 2,1. Ein prämierter Artikel wird also im Schnitt fast drei Mal häufiger zitiert als ein nicht ausgezeichneter Artikel. Die Hypothese, dass prämierten Artikeln eine größere Aufmerksamkeit in der Fachgemeinschaft zuteilwird, kann sehr klar und deutlich belegt werden (Tab. 1).

Auch unter den meistzitierten Artikeln der sozialwissenschaftlichen Fachzeitschriften, die bei ISI registriert sind und sich am Thyssen-Preis beteiligen, sind prämierte Artikel überproportional häufig vertreten (siehe Online-Anhang 2).

Besonders prägnant ist der Zusammenhang, wenn weiter nach dem ersten, zweiten und dritten Preis aufgeschlüsselt wird. Nach zehn Jahren wird ein mit dem ersten Preis

Tab. 1: Zitationshäufigkeit für die prämierten Aufsätze und für die Vergleichsstichprobe

		5 Jahre	10 Jahre
Kontrollstichprobe	N	141	113
	Mittelwert [ohne Selbstzitate]	1,12 [0,91]	2,06 [1,73]
	Standardabweichung	1,61	3,14
	Median	0	1
	Minimum/Maximum	0/8	0/16
Thyssenpreis	N	94	76
	Mittelwert [ohne Selbstzitate]	3,30 [2,81]	6,09 [5,24]
	Standardabweichung	3,29	6,93
	Median	2	4
	Minimum/Maximum	0/14	0/42

10 Jahre: Zeitraum von 10 Jahren nach Erscheinungsjahr (ohne Erscheinungsjahr)

5 Jahre: Zeitraum von 5 Jahren nach Erscheinungsjahr (ohne Erscheinungsjahr)

Einfaktorielle Varianzanalyse der Zitationshäufigkeiten für 5 Jahre **mit** Selbstziten **[ohne Selbstzitate]**: $F(1, 233)=45,59 [37,91]$; $p<0,0001 [p<0,0001]$

Einfaktorielle Varianzanalyse der Zitationshäufigkeiten für 10 Jahre **mit** Selbstziten **[ohne Selbstzitate]**: $F(1, 187)=29,35 [24,29]$; $p<0,0001 [p<0,0001]$

Mann-Whitney U-Test: Unterschiede der mittleren Zitationshäufigkeiten zwischen Kotrollgruppe und Thyssenpreis gesamt (d. h. alle drei Thyssenpreistränge zusammen): 5 Jahre **mit** Selbstziten **[ohne Selbstzitate]**: $p<0,001 [p<0,001]$; 10 Jahre **mit** Selbstziten **[ohne Selbstzitate]**: $p<0,001 [p<0,001]$

Tab. 2: Zitationshäufigkeit aller prämierten Aufsätze nach den drei Thyssenpreisen aufgeschlüsselt

		mit S.	
		5 Jahre	10 Jahre
1. Thyssenpreis	N	22	17
	Mittelwert [ohne Selbstzitate]	4,68 [4,18]	9,41 [8,47]
	Standardabweichung	3,72	6,43
	Median	3,5	8
	Minimum/Maximum	0/12	0/21
2. Thyssenpreis	N	32	26
	Mittelwert [ohne Selbstzitate]	3,47 [2,91]	6,50 [5,58]
	Standardabweichung	3,60	9,16
	Median	2	3,5
	Minimum/Maximum	0/14	0/42
3. Thyssenpreis	N	40	33
	Mittelwert [ohne Selbstzitate]	2,40 [1,98]	4,06 [3,30]
	Standardabweichung	2,46	3,99
	Median	2	3
	Minimum/Maximum	0/10	0/17

Einfaktorielle Varianzanalyse der Zitationshäufigkeiten zwischen den Gruppen der drei Thyssenpreise für 5 Jahre **mit** Selbstziten **[ohne** Selbstzitate]: $F(2, 91)=3,69$ [3,52]; $p=0,029$ [$p=0,034$]

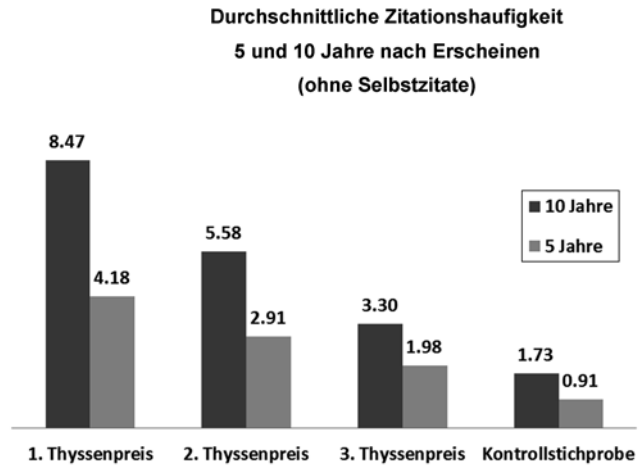
Einfaktorielle Varianzanalyse der Zitationshäufigkeiten zwischen den Gruppen der drei Thyssenpreise für Jahre 10 **mit** Selbstziten **[ohne** Selbstzitate]: $F(2, 73)=3,66$ [3,55]; $p=0,031$ [$p=0,034$]

Kruskal-Wallis Test: Unterschiede der mittleren Zitationshäufigkeiten zwischen den Thyssenpreisen (unter Ausschluss der Kontrollstichprobe) **mit** Selbstziten **[ohne** Selbstzitate]: für 5 Jahre: $p=0,040$ [$p=0,048$]; für 10 Jahre: $p=0,013$ [$p=0,016$]

prämierter Aufsatz im Durchschnitt 9,4mal zitiert, beim zweiten Preis beträgt der Wert 6,5 und beim dritten Preis sind es 4,1, während in der Vergleichsstichprobe nur zwei Zitationen registriert werden können (Tab. 2). Auch wenn Selbstzitate nicht berücksichtigt werden, bleibt der Zusammenhang erhalten (Abb. 1).

Die Zitationen im SSCI ergeben sich aus einer Art demokratischer Aufmerksamkeitsabstimmung der schreibenden Zunft. Es ist schon erstaunlich, wie stark das Expertenurteil zur Qualität der Aufsätze (im Durchschnitt) mit dieser „Abstimmung“ korreliert. Allerdings darf man auch nicht übersehen, dass die Verteilungen der Zitationen deutlich rechtsschief ausgeprägt sind und eine hohe Streuung aufweisen (Tab. 2). Es gibt also auch preisgekrönte Artikel, die selten zitiert werden und natürlich auch viele Artikel, die keinen Preis erhalten haben, dennoch aber oft zitiert werden. Trotz starker Überlappung ergibt sich aber im Durchschnitt ein klar erkennbares Muster zwischen Juryurteil und Resonanz in der Fachgemeinschaft. Das gleiche Bild zeigt sich auch anhand von Google-Scholar-Zitationen, wobei hier von Jansen (2012) zusätzlich die nominierten, aber nicht ausgezeichneten Artikel berücksichtigt wurden. Die durchschnittliche Zahl der Zitationen der nominierten Artikel lag zwischen den Zitationen für den dritten Preis und den Zitationen der nicht-nominierten Artikel, sodass sich insgesamt folgende, aufsteigende Reihung

Abb. 1: Mittelwerte der Zitationshäufigkeit für Thyssenpreise und Kontrollstichprobe



der Mittelwerte für die fünf Kategorien von Arbeiten ergab: 1) Nicht-nominierte Artikel, 2) nominierte, aber nicht preisgekrönte Arbeiten, 3) Artikel, die mit dem dritten Preis ausgezeichnet wurden, 4) Artikel, die den zweiten Preis erhielten und 5) Artikel, die mit dem ersten Thyssen-Preis ausgezeichnet wurden (Jansen 2012, S. 51, Tab. 3).

Ein Nebeneffekt unserer Arbeit ist die Generierung einer Zufallsstichprobe aus dem Korpus deutschsprachiger Fachzeitschriften. Damit gewinnen wir einen Anhaltspunkt, wie häufig ein Artikel im Durchschnitt (in ISI-Zeitschriften) zitiert wird. Das Ergebnis lautet: zwei Mal im Zeitraum von zehn Jahren nach Veröffentlichung. Ob dies im internationalen Vergleich viel oder wenig ist, könnte man feststellen, wenn für andere Länder entsprechende Zufallsstichproben gezogen werden. Üblicherweise wird die angelsächsische oder speziell US-amerikanische Dominanz beklagt, wobei aber einfach die Impact-Faktoren von Top-Zeitschriften wie „American Sociological Review“, „American Journal of Sociology“ und „Social Forces“ mit denjenigen der deutschsprachigen Zeitschriften verglichen werden (z. B. Münch 2009). Dieser Vergleich ist aber fragwürdig. Nur ein extrem kleiner Anteil der amerikanischen Sozialwissenschaftler publiziert in den Spitzenzeitschriften, deren Autorenschaft sich zudem international rekrutiert. Schließlich existieren in den USA auch eine große Zahl allgemeiner und spezieller soziologischer Fachzeitschriften, die in einem unverzerrten Vergleich berücksichtigt werden müssten. Vergleicht man die Rezeption der Artikel amerikanischer Soziologen auf einer entsprechend breiteren Basis von Fachzeitschriften, könnte sich durchaus herausstellen, dass ein englischsprachiger Artikel eines amerikanischen Autors im Mittel nicht häufiger rezipiert wird als ein Artikel in den deutschsprachigen Fachzeitschriften.

5 Soziale Konstruktion von Aufmerksamkeit?

Es gibt anekdotische Beispiele dafür, dass Autoren ihre eigenen Bücher gekauft haben, um sie auf Bestsellerlisten zu platzieren. Das Etikett „Bestseller“ wiederum kann eine Dynamik wachsender Verkaufszahlen in Gang bringen (Keuschnigg 2012). Wir gehen im Folgenden der Frage nach, ob nicht die Qualität der Artikel an sich, sondern erst das von

Experten zugewiesene Gütesiegel zu erhöhter Aufmerksamkeit in der Fachgemeinschaft verholfen hat. Der Indikator „Zitationen“ wäre dann kein externes Kriterium für die Validität des Expertenurteils. Vielmehr würden die Experten erst durch ihre Zuschreibungen soziale Realität schaffen. Ihr Urteil würde sich dann im Sinne einer sich selbst erfüllenden Prognose nachträglich quasi zwangsläufig bestätigen, unabhängig von den Eigenschaften eines prämierten Artikels.

Wie ließe sich die These überprüfen? Wir könnten die Jury bitten, die Preise bei der nächsten Runde der Preisverleihung per Zufall zu verteilen und dies geheim zu halten. Einige Jahre später könnten wir (nahezu) zweifelsfrei herausfinden, ob die These einer durch Definitionsprozesse selbsterzeugten Realität zutreffend ist. Aus guten Gründen haben wir uns für eine andere Methode entschieden. Die Jury bestimmt die Preisträger erst im Jahr nach Erscheinen der prämierten Artikel. Zwischen dem Publikationsdatum und der Bekanntgabe der Preisträger vergeht oft mehr als ein Jahr. Eine weitere Frist kommt hinzu, da es eine Weile dauert, bis die Zitation in ein neues Manuskript einfließt und dieses veröffentlicht wird. Wir sind aber auf der sicheren Seite, wenn wir von einer Latenzzeit von durchschnittlich einem Jahr ausgehen. Gibt es keinen eigenständigen Rezeptionseffekt aufgrund (im Durchschnitt) genuin hervorstechender Eigenschaften der Publikationen, dann dürften sich die Zitationsraten der prämierten Arbeiten in dieser Periode nicht von den Zitationsraten der Artikel in der Vergleichsstichprobe unterscheiden.

Die beobachteten Zitationsraten legen hingegen einen anderen Schluss nahe. Auch bereits vor einem möglichen Effekt der Zuschreibung divergieren die Zitationsraten zwischen preisgekrönten und anderen Artikeln. Der Mittelwert der Thyssen-Aufsätze beträgt 0,76, der Mittelwert in der Vergleichsstichprobe ist 0,18. Erwartungsgemäß sind die Zahlen gering, aber die Differenz ist deutlich erkennbar. Mehr noch: Auch die perfekte Übereinstimmung der Rangfolge der Zitationsraten mit der Rangfolge der Preise zeigt sich bereits nach der kurzen Periode von etwa einem Jahr. Die Zitationsrate für den ersten Preis beträgt 1,09, für den zweiten Preis sind es 0,78 und beim dritten Preis ist der Wert 0,58 (Tab. 3). Die Artikel unterscheiden sich also schon, bevor die Preisverleihung eine Wirkung entfaltet haben kann. Nun könnte eingewendet werden, die Unterschiede seien darauf zurückzuführen, dass mit dem Thyssen-Preis prämierte Autoren womöglich häufiger publizieren und sich entsprechend häufiger zitieren als andere Autoren.⁶ Der Einwand sticht allerdings nicht. Denn unsere Auswertung ohne Selbstzitate zeigt, dass der Effekt robust ist. Ohne Selbstzitate sind es 0,87 für den ersten, 0,53 für den zweiten, 0,33 für den dritten Preis und 0,14 Zitationen in der Kontrollgruppe (Tab. 3).

Wir haben mit unserer Analyse gezeigt, dass die Jury Artikel auswählt, die sich durch eine in den Folgejahren erhöhte Rezeption auszeichnen. Damit ist aber noch nicht nachgewiesen, dass das konstruktivistische Moment der „Etikettierung“ keinen Effekt ausübt. Es könnte beides der Fall sein. Ausgezeichnete Artikel haben Eigenschaften, die zu einem höheren „impact“ führen; dieser wird aber zusätzlich durch die Prämierung verstärkt. Wir können diese Frage zu beantworten versuchen, indem wir die Stärke der Unterschiede im

6 Wir hatten ursprünglich Selbstzitate nicht ausgeschlossen. Wir bedanken uns für den Hinweis auf mögliche systematische Verzerrungen durch Selbstzitate und die Anregung zur graphischen Untersuchung der Verlaufskurven bei der anonymen Gutachterin und dem Gutachter.

Tab. 3: Zitationshäufigkeit für die prämierten Aufsätze und für die Vergleichsstichprobe

		1 Jahr	
Kontrollstichprobe	N	146	
	Mittelwert [ohne Selbstzitate]	0,18 [0,14]	
	Standardabweichung	0,53	
	Median	0	
	Minimum/Maximum	0/3	
1. Thyssenpreis	N	23	<i>Thyssenpreis gesamt</i>
	Mittelwert [ohne Selbstzitate]	1,09 [0,87]	
	Standardabweichung	1,20	
	Median	1	
	Minimum/Maximum	0/5	
2. Thyssenpreis	N	36	102
	Mittelwert [ohne Selbstzitate]	0,78 [0,53]	0,76 [0,52]
	Standardabweichung	1,07	1,03
	Median	0	0
	Minimum/Maximum	0/4	0/5
3. Thyssenpreis	N	43	
	Mittelwert [ohne Selbstzitate]	0,58 [0,33]	
	Standardabweichung	0,85	
	Median	0	
	Minimum/Maximum	0/3	

1 Jahr: Erscheinungsjahr und das darauf folgende Jahr

Einfaktorielle Varianzanalyse der mittleren Zitationshäufigkeiten von Kontrollstichprobe vs. Thyssenpreis gesamt für 1 Jahr **mit** Selbstziten [ohne Selbstzitate]: $F(1, 246)=33,93$ [23,84]; $p<0,0001$ [$p<0,0001$]

Einfaktorielle Varianzanalyse der mittleren Zitationshäufigkeiten von Kontrollstichprobe und den Gruppen der drei Thyssenpreise für 1 Jahr **mit** Selbstziten [ohne Selbstzitate]: $F(3, 244)=13,71$ [12,47]; $p<0,0001$ [$p<0,0001$]

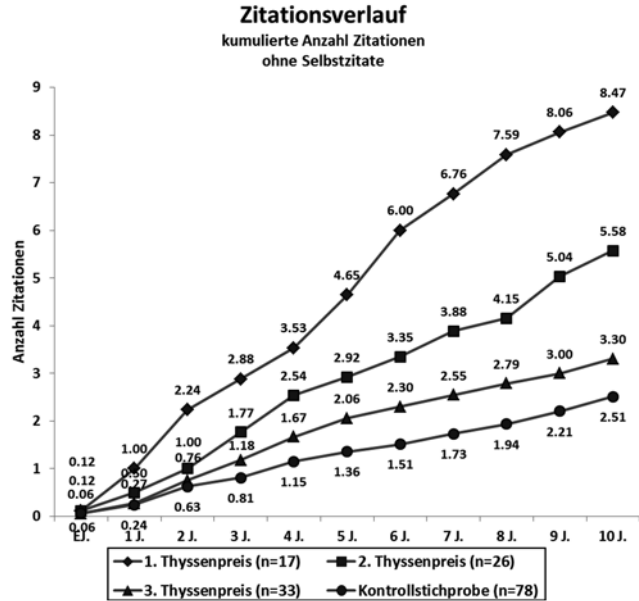
Einfaktorielle Varianzanalyse der mittleren Zitationshäufigkeiten der drei Thyssenpreise (unter Ausschluss der Kontrollgruppe) für 1 Jahr **mit** Selbstziten [ohne Selbstzitate]: $F(2, 99)=1,85$ [3,60]; $p=0,162$ [$p=0,0310$]

Mann-Whitney U-Test: Unterschied der Zitationshäufigkeiten zwischen Kotrollgruppe und Thyssenpreis gesamt für 1 Jahr **mit** Selbstziten [ohne Selbstzitate]: $p<0,001$ [$p<0,001$]

Kruskall-Wallis Test: Unterschiede der Zitationshäufigkeiten von Kontrollstichprobe und den Gruppen der drei Thyssenpreise für 1 Jahr **mit** Selbstziten [ohne Selbstzitate]: $p<0,001$ [$p<0,001$]

Kruskall-Wallis Test: Unterschiede zwischen den drei Thyssenpreisen (unter Ausschluss der Kontrollgruppe) für 1 Jahr **mit** Selbstziten [ohne Selbstzitate]: $p=0,149$ [$p=0,042$]

Abb. 2: Kumulierte durchschnittliche Zitationshäufigkeiten der drei Thyssenpreise und der Kontrollstichprobe über zehn Jahre. In diese Grafik sind nur Fälle eingegangen, für die im Beobachtungszeitraum von zehn Jahren Daten vorliegen



ersten Jahr mit den Differenzen nach fünf und zehn Jahren vergleichen. Im ersten Jahr nach Erscheinen werden Artikel, die den ersten Preis gewonnen hatten, 1,09 Mal zitiert, nicht-prämierte Arbeiten 0,18 Mal. Nach fünf Jahren sind es 4,68 versus 1,12 und für die Zeitspanne von zehn Jahren lauten die entsprechenden Zahlen 9,41 versus 2,06. Im ersten Jahr werden die höchst prämierten Artikel demnach rund sechs Mal öfter zitiert, nach fünf Jahren beträgt der entsprechende Faktor nur mehr 4,2 und nach zehn Jahren 4,6. Ähnlich verhält es sich, wenn die Selbstzitate weggelassen werden. Die Relationen betragen nach einem Jahr 6,2, nach fünf Jahren 4,6 und nach zehn Jahren 4,9. Die Relation wächst demnach nach der Bekanntgabe des Preises nicht. Wir haben zudem die kumulativen Verlaufskurven für Artikel ermittelt, für die Beobachtungen über eine Zeitspanne von zehn Jahren vorlagen. Die Fallzahlen in den einzelnen Preisgruppen sind dann allerdings relativ gering. Wieder unterscheiden sich die einzelnen Kurven entsprechend der Preisrangfolge bis hin zur Kontrollgruppe deutlich voneinander. Betrachtet man die Verlaufskurven in Abb. 2 genauer, dann findet man zwar beim ersten und dritten Preis einen überdurchschnittlichen Anstieg vom ersten zum zweiten Jahr. Allerdings ist der Anstieg vom fünften auf das sechste Jahr für den ersten Preis noch ausgeprägter und in der Kontrollgruppe liegt der Zuwachs vom ersten auf das zweite Jahr ebenfalls über dem Durchschnitt. Insgesamt lässt sich vermuten, dass die stärkere Rezeption vor allem durch von Anfang an gegebene Eigenschaften der ausgezeichneten Arbeiten hervorgerufen wird.

6 Ausblick

Wir haben in dieser Arbeit einen deutlichen und signifikanten Zusammenhang zwischen dem Jury-Urteil über die Vergabe des Thyssen-Preises und der Zitationsrate der prämier-

ten Artikel belegen können. Dabei wurden die Ergebnisse für die prämierten Artikel mit einer Kontrollstichprobe der nicht ausgezeichneten Arbeiten verglichen. Prämierte Artikel „ernten“ signifikant mehr Zitationen als die „normalen“ Artikel. Bemerkenswert ist, dass das Muster der Zitationsrangfolge exakt der Rangfolge der Preise entspricht. Der erste Preis erhält die meisten, der zweite Preis die zweitmeisten und der dritte Preis die drittmeisten Zitationen, während nicht prämierte Artikel an vierter Stelle der Rangfolge stehen. Der Unterschied zeigt sich auch, wenn nur Zitationen im Jahr der Veröffentlichung eines Artikels und im Folgejahr betrachtet werden. Mehr noch: Auch im ersten Jahr nach der Publikation zeigt sich bereits exakt die gleiche Reihenfolge der durchschnittlichen Zitationshäufigkeiten. In diesem Zeitraum kann der Preis „an sich“ noch keine Wirkung auf die Aufmerksamkeit entfaltet haben. Die Jury hat demnach ein Gespür, Artikel auszuzeichnen, die tatsächlich und unabhängig von der Preisvergabe in der „Scientific Community“ erhöhte Aufmerksamkeit reklamieren können.

Unsere empirische Analyse ergibt keine Anhaltspunkte für die These, dass der Preis an sich eine Neudefinition der Situation hervorruft und damit scheinbare Qualität in der Wahrnehmung der Fachöffentlichkeit erzeugt. Nicht immer gilt, dass des Kaisers neue Kleider imaginär sind. Wir behaupten allerdings nicht, dass konstruktivistische Elemente bei der Bewertung und Verbreitung wissenschaftlicher Erkenntnisse keine Rolle spielten. Nur gilt die These nicht generell und vor allem kann nicht einfach vom Schreibtisch des Gelehrten aus dekretiert werden, ob und in welchem Ausmaß sozialkonstruktivistische Hypothesen zutreffend sind oder eben nicht. Das Thomas-Theorem ist kein universelles Gesetz, mit dem sich prognostizieren lässt, welche einer Situation zugeschriebenen Eigenschaften bei welchen Akteuren bestimmte Verhaltensweisen hervorrufen. Wichtig für die Forschung wäre herauszufinden, unter *welchen Bedingungen* Etikettierungsprozesse beobachtbares Verhalten beeinflussen.⁷

Hinzu kommt, dass Reputation und Signale im Wettbewerb um Aufmerksamkeit angesichts knapper Zeit-Ressourcen eine durchaus rationale Filterwirkung ausüben können. Der Preis dafür ist „statistische Diskriminierung“, denn das Renommee einer Zeitschrift z. B. bezieht sich nur auf den Durchschnitt und lässt keinesfalls einen eindeutigen Schluss auf die Qualität einer Arbeit zu. Das Wissenschaftssystem ist allerdings nicht schon deshalb irrational, wenn Reputation und Signale die Aufmerksamkeit auf bestimmte, ausgewählte Erkenntnisse lenken. Ein Systemversagen würde erst dann vorliegen, wenn dadurch ein systematischer „Bias“ erzeugt wird. Bei anderen, hier nicht zur Sprache gekommenen Elementen des Wissenschaftssystems gibt es dafür sehr deutliche Anhaltspunkte. So führt die überdurchschnittlich häufige Publikation signifikanter und bestätigender Resultate ebenso zu einer systematischen Verzerrung wie eine hohe Fehlerquote falsch-positiver Resultate, die sich aus der Logik von Signifikanztests ergibt (Ioannidis 2005).⁸

7 Das Ziel des Thyssenpreises ist ja auch weder in der Begünstigung des „Zitationsschicksals“ der ausgezeichneten Artikel noch im Setzen *neuer* Maßstäbe für wissenschaftliche Qualität, sondern in der Würdigung herausragender wissenschaftlicher Leistungen zu sehen. „Soziale Tatsachen schaffende Wirkungen“ des Thyssenpreises im Sinne des Sozialkonstruktivismus müssten eher im Bereich akademischer Berufungsverfahren gesucht werden.

8 Siehe zu diesem Thema das Sonderheft der Zeitschrift für Nationalökonomie und Statistik über Artefakte (Diekmann 2011). Ein Gutachten hat kritisiert, dass die hier berichtete Forschung

Aus unseren Ergebnissen folgt nicht der Nachweis der Preisvergabe nach universalistischen Kriterien im Sinne von Merton (1973). Es wäre durchaus möglich, dass die Jury Artikel mit bestimmten Themen, verfasst von bekannten Autoren an prestigeträchtigen Institutionen bevorzugt. Da die Jury-Mitglieder mehrheitlich die Fachrichtung Soziologie vertreten, könnte man auch vermuten, dass soziologische Artikel eher als Arbeiten aus anderen Fachrichtungen als preiswürdig eingeschätzt werden (Alber und Fliegner 2010). Gleichzeitig könnten diese und andere Merkmale auch mit einer höheren Zitationsrate in der Fachgemeinschaft einhergehen. Auch Artikel, die in den häufiger zitierten Fachzeitschriften veröffentlicht wurden, könnten bei der Preisvergabe wissentlich oder unwissentlich mit einem Bonus bedacht worden sein (Jansen 2012). Ob und in welchem Ausmaß diese und weitere partikularistische Kriterien eine Rolle bei der Prämierung spielen, können wir anhand unserer Daten nicht beurteilen. Einige Hinweise auf Zusammenhänge dieser Merkmale mit der Jury-Entscheidung findet man in den Studien von Alber und Fliegner (2010), Alber et al. (2010) sowie Jansen (2012). Aber selbst wenn man in Regressionsanalysen herausfindet, dass z.B. die Bekanntheit eines Autors die Wahrscheinlichkeit einer Prämierung erhöht, folgt daraus keineswegs zwingend, dass nach partikularistischen Kriterien entschieden wurde. Denn es ist ja nicht völlig unplausibel, dass erfolgreiche Autoren (im Schnitt) bessere Qualität liefern als weniger erfolgreiche Autoren. Gleiches gilt für das Prestige eines Instituts, das Renommee einer Fachzeitschrift oder andere Merkmale, die mit der Qualität der Forschung vermutlich korreliert sind.

Es sollte aufgefallen sein, dass wir generell Aussagen über die wissenschaftliche Qualität von Artikeln vermieden haben. So lange es, unabhängig vom Jury-Urteil, keinen Konsens über valide Indikatoren der Qualität wissenschaftlicher Artikel gibt, wird man auch die Frage nach dem Einfluss partikularistischer oder universalistischer Kriterien bei den Jury-Entscheidungen nicht zweifelsfrei beantworten können.

Danksagung: Für wertvolle Anregungen bedanken wir uns bei den anonymen Gutachterinnen und Gutachtern und den Mitgliedern der Redaktion. Die Daten stehen zur Nachprüfung und für Sekundäranalysen zur Verfügung. Der Datensatz und die Liste der Artikel in den Stichproben (Online-Anhang) sind unter der folgenden Webadresse zu finden: <http://www.socio.ethz.ch/research/datafiles>.

Literatur

Alber, Jens, und Florian Fliegner, 2010. Rezeption und Themen der mit dem Preis der Fritz-Thyssen-Stiftung ausgezeichneten sozialwissenschaftlichen Aufsätze. *Soziologie* 39:286–309.

„affirmativ“ sei und besser Artikel veröffentlicht werden sollten, die Fehlentwicklungen im Wissenschaftssystem benennen. Abgesehen davon, dass der Erstautor dieses Artikels an verschiedenen Stellen systematische und institutionelle oder durch methodische Regeln hervorgerufene Fehlentwicklungen aufgezeigt hat (zuletzt Diekmann 2011), wäre die Konsequenz dieser Empfehlung in höchstem Maße fragwürdig. Angenommen in der Evaluationsforschung würden nur „nicht-affirmative“ Ergebnisse publiziert, z.B. in der Wissenschafts-, Bildungs- und Entwicklungsforschung nur Ergebnisse über Fehlentwicklungen. Manuskripte mit anderen Ergebnissen würden von „nicht-affirmativen Gutachtern“ zur Ablehnung der Publikation empfohlen. Die Folge wäre ein systematischer Bias in der veröffentlichten Forschung.

- Alber, Jens, Florian Fliegner und Torben Nerlich. 2010. Was bestimmt Exzellenz in den Sozialwissenschaften? Eine Analyse am Beispiel des Preises der Fritz-Thyssen-Stiftung für sozialwissenschaftliche Aufsätze 1981–2006. *Soziologie* 39:152–178.
- Baldi, Stéphane. 1998. Normative versus social constructivist processes in the allocation of citations: A network-analytic model. *American Sociological Review* 63:829–846.
- Bornmann, Lutz, und Hans-Dieter Daniel. 2006. Selecting scientific excellence through committee peer review. A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowships applicants. *Scientometrics* 68:427–440.
- Bornmann, Lutz, und Hans-Dieter Daniel. 2008. What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation* 64:45–80.
- Diekmann, Andreas, (Hrsg.). 2011. *Methodological artefacts, data manipulation and fraud in economics and social science*. Stuttgart: Lucius & Lucius.
- Gottfredson, Stephen D. 1978. Evaluating psychological research reports. Dimensions, reliability, and correlates of quality judgments. *American Psychologist* 33:920–934.
- Haslam, Nick, Lauren Ban, Leah Kaufmann, Stephen Loughnan, Kim Peters, Jennifer Whelan und Sam Wilson. 2008. What makes an article influential? Predicting impact in social and personality psychology. *Scientometrics* 76:169–185.
- Ioannidis, John P. A. 2005. Why most published research findings are false. *PLoS Med* 2:696–701.
- Jansen, Marcus. 2010. *Daten zu Merkmalen von Zeitschriftenpublikationen in der deutschsprachigen Sozialwissenschaft zwischen den Jahren 1981 und 2006*. Unveröffentlichtes Manuskript. Köln: Forschungsinstitut für Soziologie der Universität zu Köln.
- Jansen, Marcus. 2012. Reputationsdimensionen in den Sozialwissenschaften. *Soziologie* 41:28–55.
- Judge, Timothy A., Daniel M. Cable, Amy E. Colbert und Sara L. Rynes. 2007. What causes a management article to be cited – article, author, or journal? *Academy of Management Journal* 50:491–506.
- Keuschnigg, Marc. 2012. *Das Bestseller-Phänomen. Die Entstehung von Nachfragekonzentration im Buchmarkt*. Wiesbaden: VS-Verlag für Sozialwissenschaften.
- Knorr-Cetina, Karin. 1989. Spielarten des Konstruktivismus. *Soziale Welt* 40:86–96.
- Lee, John D., Kim J. Vicente, Andrea Cassano und Anna Shearer. 2003. Can scientific impact be judged prospectively? A bibliometric test of Simonton's model of creative productivity. *Scientometrics* 56:223–233.
- Merton, Robert K. 1968. The Matthew effect in science. *Science* 159:56–63.
- Merton, Robert K. 1973. *The sociology of science. Theoretical and empirical investigations*. Chicago: University of Chicago Press.
- Merton, Robert K. 1995. The Thomas theorem and the Matthew effect. *Social Forces* 74:379–424.
- Münch, Richard. 2009. Publikationsverhalten in der Soziologie. In *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen*, 2. erw. Aufl., Hrsg. Alexander von Humboldt Stiftung, 69–77. Bonn.
- Neidhardt, Friedhelm. 2008. Das Forschungsrating des Wissenschaftsrats. Einige Erfahrungen und Befunde. *Soziologie* 37:421–432.
- Sternberg, Robert J., und Tamara Gordeeva. 1996. The anatomy of impact: What makes an article influential? *Psychological Science* 7:69–75.
- Thomas, William I., und Dorothy S. Thomas. 1928. *The child in America: Behavior problems and programs*. New York: Knopf.

Andreas Diekmann, Dr., Professor für Soziologie an der ETH-Zürich. Forschungsschwerpunkte: Soziale Kooperation, Umweltsoziologie, Methoden der Sozialforschung. Neuere Veröffentlichungen: Die Rolle sozialer Normen, der Situationsdefinition und sozialer Klassen beim Untergang der Titanic. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 64, 2012. Temporal Embeddedness and Signals of Trustworthiness: Experimental Tests of a Game Theoretical Model in the UK, Russia and Switzerland. Erscheint in *European Sociological Review* (mit Wojtek Przepiorka). Making Use of „Benford’s Law“ for the Randomized Response Technique. Erscheint in *Sociological Methods and Research* 41, 2012.

Matthias Näf, Wissenschaftliche Hilfskraft am Lehrstuhl für Soziologie, ETH-Zürich, Schwerpunkt Umweltsoziologie.

Manuel Schubiger, Student der Informationstechnologie und Elektrotechnik an der ETH Zürich.